AFRL-AFOSR-VA-TR-2021-0105

Complementing dynamical equations with data in adaptive reduced-order subspaces

Sapsis, Themistoklis
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA, 2139
USA

08/17/2021
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
17-08-2021	Final	15 Jun 2016 - 14 Jun 2019

4. TITLE AND SUBTITLE
Complementing dynamical equations with data in adaptive reduced-order subspaces

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-16-1-0231

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Themistoklis Sapsis

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA 2139
USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR RTA2

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-VA-TR-2021-0105

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This was a no-cost extension with limited annual budget, corresponding to 3 months of postdoc support. Overt this period we worked on finalizing several manuscripts on optimal sampling for the prediction and quantification of extreme events. A review article on extremes in fluids and waves was submitted and published to the Annual Review of Fluid Mechanics, during this period.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			FARIBA FAHROO
U	U	U	UU	11	**19b. TELEPHONE NUMBER** *(Include area code)* 426-8429

Standard Form 298 (Rev.8/98)
Prescribed by ANSI Std. Z39.18

Final report: 6/15/2016-6/14/2020

Complementing dynamical equations with data in adaptive reduced-order subspaces

Project # FA9550-16-1-0231

PI: Themistoklis Sapsis, MIT

1 Objectives

Leveraging the power of the optimally time-dependent (OTD) modes, a set of deformable orthonormal tangent vectors that track directions of instabilities along any trajectory, we wish to:

1. Develop mathematical/computational algorithms for the identification of precursors for rare extreme events using adaptive reduced order methods. We plan to employ both the governing equations, if/when available, but also active learning methods specially designed for extreme transient events.

2. Apply these algorithms to prototype dynamical systems excibiting extreme rare events encountered in engineering and nature.

3. Design reduced-order control algorithms capable of predicting and suppressing transient (non-normal) and asymptotic growth around any fixed point of the governing equations;

4. Demonstrate robustness of the developed control methods with respect to the perturbation amplitude, and formulate a localized control strategy in which actuation is realized in a subdomain of the physical domain of interest; and

5. Exploit the versatility of neural networks to learn the "pointwise" mapping from phase space to OTD space directly from data, resulting in a cartography of directions associated with strongest instabilities in phase space.

2 Publications and presentations

This effort resulted in 25 publications in Applied Math and Engineering journals. The following items were produced through the support of this effort (in chronological order):

Journal publications

- H. Babaee, T. Sapsis, A minimization principle for the description of time-dependent modes associated with transient instabilities, Proceedings of the Royal Society A, 472 (2016) 20150779. ***[Featured in the journal's cover page]***.
- M. Farazmand, T. Sapsis, Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems, Physical Review E, 94 (2016) 032212 (15 pages). ***[Featured in the journal's kaleidoscope]***.
- Z. Y. Wan, T. Sapsis, Reduced-space Gaussian process regression for data-driven probabilistic forecast of chaotic dynamical systems, Physica D, 345 (2017) 40-55.
- M. Farazmand, T. Sapsis, Reduced-order prediction of rogue waves in two dimensional water waves, Journal of Computational Physics, 340 (2017) 418-434.
- H. Babaee, M. Farazmand, G. Haller, T. Sapsis, Reduced-order description of transient instabilities and computation of finite-time Lyapunov exponents, Chaos, 27 (2017) 063103.
- H. -K. Joo, M. Mohamad, T. Sapsis, Extreme events and their optimal mitigation in nonlinear structural systems excited by stochastic loads: Application to ocean engineering systems, Ocean Engineering Journal, 142 (2017) 145-160.
- M. Farazmand, T. Sapsis, A variational approach to probing extreme events in turbulent dynamical systems, **Science Advances**, 3:e1701533 (2017).
- H. -K. Joo, M. Mohamad, T. Sapsis, Heavy-tailed response of structural systems subjected to extreme forcing events, ASME Journal of Computational and Nonlinear Dynamics, 13 (2018) 090914 (12 pages).
- M. Farazmand, T. Sapsis, Physics-based probing and prediction of extreme events, **SIAM News**, 51 (2018) 1.
- S. Mowlavi, T. Sapsis, Model order reduction for stochastic dynamical systems with continuous symmetries, SIAM Journal on Scientific Computing, 40 (2018) 1669-1695.
- P. Vlachas, W. Byeon, Z. Y. Wan, T. Sapsis, P. Koumoutsakos, Data-driven forecasting of high-dimensional chaotic systems with long-short term memory networks, **Proceedings of the Royal Society A**, 474 (2018) 20170844 (20 pages).
- Z. Y. Wan, P. Vlachas, P. Koumoutsakos, T. Sapsis, Data-assisted reduced-order modeling of extreme events in complex dynamical systems, PLOS One, 24 May (2018) (22 pages).
- M. Mohamad, T. Sapsis, A sequential sampling strategy for extreme event statistics in nonlinear dynamical systems, **Proceedings of the National Academy of Sciences**, 115 (2018) 11138-11143.

- A. Blanchard, S. Mowlavi, T. Sapsis, Control of linear instabilities by dynamically consistent order reduction on optimally time-dependent modes, Nonlinear Dynamics, 95 (2019) 2745-2764.
- P. Blonigan, M. Farazmand, T. Sapsis, Are extreme dissipation events predictable in turbulent fluid flows?, Physical Review Fluids, 4 (2019) 044606 (21 pages).
- A. Blanchard, T. Sapsis, Analytical description of optimally time-dependent modes for reduced-order modeling of transient instabilities, SIAM Journal on Applied Dynamical Systems, 18 (2019) 1143-1162.
- A. Blanchard, T. Sapsis, Stabilization of unsteady flows by reduced-order control with optimally time-dependent modes, Physical Review Fluids, 4 (2019) 053902.
- M. Farazmand, T. Sapsis, Closed-loop adaptive control of extreme events in a turbulent flow, Physical Review E, 100 (2019) 033110 (7 pages).
- S. Guth, T. Sapsis, Machine learning predictors of extreme events occurring in complex dynamical systems, Entropy, 21 (2019) 925 (18 pages).
- A. Blanchard, T. Sapsis, Learning the tangent space of dynamical instabilities from data, Chaos, 29 (2019) 113120, Focus Issue: When Machine Learning Meets Complex Systems: Networks, Chaos and Nonlinear Dynamics, (15 pages).
- Z. Vlachas, J. Pathak, B. R. Hunt, T. Sapsis, M. Girvan, E. Ott, P. Koumoutsakos, Forecasting of Spatio-temporal Chaotic Dynamics with Recurrent Neural Networks: a comparative study of Reservoir Computing and Backpropagation Algorithms, Neural Networks, 126 (2020) 191-217.
- T. Sapsis, Output-weighted optimal sampling for Bayesian regression and rare event statistics using few samples, **Proceedings of the Royal Society A**, 476 (2020) 20190834 (24 pages).

In addition the following invited review articles were published:

Invited Review Articles

- T. Sapsis, New perspectives for the prediction and statistical quantification of extreme events in high-dimensional dynamical systems, **Philosophical Transactions of the Royal Society A**, 376 (2018) 20170133 (18 pages).
- M. Farazmand, T. Sapsis, Extreme events: mechanisms and prediction, ASME Applied Mechanics Reviews, 71 (2019) 050801. **[Lloyd Hamilton Donnell Applied Mechanics Reviews Paper Award].**
- T. Sapsis, Statistics of extreme events in fluid flows and waves, **Annual Review of Fluid Mechanics**, 53 (2021) 85-111.

2.1 News Coverage

The results of this work have been featured in numerous media outlets including **The Economist**, **BBC**, **Popular Science** and others.

3 Personnel

This project partially supported 3 postdocs and 4 graduate students. Two of the three postdocs, Hessam Bababe (now in Pittsburg U.) and Mohammad Farazmand (now in NCSU), currently hold tenured track faculty positions in Mechanical Engineering and Applied Math Departments. All of the involved graduate students graduated with a PhD and currently hold positions as postdocs (NYU-Courant) or in industry (Facebook, Mathworks, Goldman Sachs).

4 Methods

Although the project supported several aspects of research related to extreme transient events in complex systems (see Publications section), here we give an overview of its main focus related to the analysis of extreme events using adaptive subspaces.

4.1 Review of the OTD modes

We consider a finite-dimensional autonomous system $\dot{\mathbf{z}} = \mathbf{F}(\mathbf{z})$, where $\mathbf{z}(t) \in \mathbb{R}^d$ is the state vector at time t, $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth vector field, and overdot denotes differentiation with respect to time. We assume that the system admits at least one fixed point $\mathbf{z}_e$. Trajectories initiated in the vicinity of $\mathbf{z}_e$ are rapidly expelled from it and ultimately settle into a different attractor $\mathscr{A}$. Infinitesimal perturbations about a given trajectory obey the variational equation

$$\dot{\mathbf{v}}_i = \mathbf{L}(\mathbf{z})\mathbf{v}_i, \quad 1 \le i \le r, \tag{1}$$

where $\mathbf{L}(\mathbf{z}) = \nabla\mathbf{F}(\mathbf{z}) \in \mathbb{R}^{d\times d}$ is the Jacobian matrix of $\mathbf{F}$ evaluated at $\mathbf{z}$. We emphasize that $\mathbf{L}(\mathbf{z})$ is a time-dependent operator because it depends on the current state $\mathbf{z}(t)$. In principle, the variational equation could be used to track directions of instabilities around trajectories. In practice, however, this is impossible, because any collection of perturbations $\{\mathbf{v}_i(t)\}_{i=1}^r$ evolved with (1) for a sufficiently long time would see the magnitude of its individual members grow or decay exponentially fast, and the angle between them rapidly vanish.

To compute a set of meaningful directions (or "modes") from the variational equation, Babaee & Sapsis [2] proposed to enforce orthonormality of the $\mathbf{v}_i(t)$'s at all times. One way to achieve this is to continuously apply the Gram–Schmidt algorithm to the collection $\{\mathbf{v}_i(t)\}_{i=1}^r$, starting with $\mathbf{v}_1$ and moving down. Blanchard & Sapsis [3] showed that the resulting vectors obey

$$\dot{\mathbf{u}}_i = \mathbf{L}(\mathbf{z})\mathbf{u}_i - \langle \mathbf{L}(\mathbf{z})\mathbf{u}_i, \mathbf{u}_i\rangle \mathbf{u}_i - \sum_{k=1}^{i-1} \left(\langle \mathbf{L}(\mathbf{z})\mathbf{u}_i, \mathbf{u}_k\rangle + \langle \mathbf{L}(\mathbf{z})\mathbf{u}_k, \mathbf{u}_i\rangle\right) \mathbf{u}_k, \quad i \in \{1, \dots, r\}. \tag{2}$$

where the angle brackets denote a suitable inner product. In the above, we recognize the variational equation (the left-hand side and the first term on the right-hand side), appended with terms enforcing the orthonormality constraint (the last two terms on the right-hand side). The $\mathbf{u}_i$'s have been referred to as the OTD modes, and the collection $\{\mathbf{u}_i\}_{i=1}^r$ as the OTD subspace [2]. We note that in matrix form, (2) may be written as

$$\dot{\mathbf{U}} = \mathbf{L}(\mathbf{z})\mathbf{U} - \mathbf{U}(\mathbf{U}^\mathsf{T}\mathbf{L}(\mathbf{z})\mathbf{U} + \mathbf{\Phi}) = \mathbf{L}(\mathbf{z})\mathbf{U} - \mathbf{U}\mathbf{L}_r(\mathbf{z}), \tag{3}$$

where $\mathbf{\Phi}$ is a skew-symmetric tensor.

A key property of the OTD modes is that they and the $\mathbf{v}_i$'s span the same subspace, so the OTD modes provide a numerically stable way of tracking instabilities. For hyperbolic fixed points, the OTD subspace aligns with the most unstable eigenspace of the associated linearized operator [2]. For generic trajectories, the OTD subspace aligns with the left eigenspace of the Cauchy–Green tensor, which characterizes transient instabilities [1]. It is thus natural to incorporate the OTD modes in a reduced-order control algorithm.

4.2 Control of instabilities by OTD modes

To formulate an OTD-based control law, we consider the controlled dynamics of an infinitesimal perturbation $\mathbf{z}' \in \mathbb{R}^d$ around a fixed point $\mathbf{z}_e$, described by $\dot{\mathbf{z}}' = \mathbf{L}\mathbf{z}' + \mathbf{B}\mathbf{c}$, where $\mathbf{L} \equiv \mathbf{L}(\mathbf{z})$ is used as a proxy for $\mathbf{L}(\mathbf{z}_e)$. Introducing the OTD projection $\boldsymbol{\eta} = \mathbf{U}^\mathsf{T}\mathbf{z}'$, and defining a reduced control matrix $\mathbf{B}_r = \mathbf{U}^\mathsf{T}\mathbf{B} \in \mathbb{R}^{r\times p}$, we obtain the reduced controlled variational equation

$$\dot{\boldsymbol{\eta}} = \mathbf{L}_r\boldsymbol{\eta} + \mathbf{B}_r\mathbf{c}, \tag{4}$$

where $\mathbf{L}_r$ is the reduced linear operator. If the control vector is sought in the form $\mathbf{c} = \mathbf{K}_r\boldsymbol{\eta}$, with $\mathbf{K}_r \in \mathbb{R}^{p\times r}$ a reduced feedback gain matrix, then (4) reduces to $\dot{\boldsymbol{\eta}} = \mathbf{L}_{r,c}\boldsymbol{\eta}$, where $\mathbf{L}_{r,c} = \mathbf{L}_r + \mathbf{B}_r\mathbf{K}_r$ is the closed-loop reduced linear operator. The latter is time-dependent, so its eigenvalues are not good indicators for growth or decay of $\|\boldsymbol{\eta}\|$. However, the eigenvalues of its symmetric part characterize the instantaneous rate of change of the magnitude of the reduced perturbation, since

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\eta}\|^2 = \frac{\langle \mathbf{L}_{r,c}\boldsymbol{\eta}, \boldsymbol{\eta}\rangle + \langle \boldsymbol{\eta}, \mathbf{L}_{r,c}\boldsymbol{\eta}\rangle}{2}. \tag{5}$$

To stabilize the fixed point $\mathbf{z}_e$, we require that the magnitude of reduced perturbations always decay (i.e., $\mathrm{d}\|\boldsymbol{\eta}\|^2/\mathrm{d}t < 0$ for all $\boldsymbol{\eta} \neq \mathbf{0}$) and, hence, that $(\mathbf{L}_{r,c} + \mathbf{L}_{r,c}^\mathsf{T})/2$ be negative definite. We note, however, that there is no general framework in control theory addressing the issue of pole placement for the *symmetric part* of a linear operator. So we make one additional assumption, namely, that the controller can act on every state of the system (i.e., $\mathbf{B} = \mathbf{I}$) and, invoking dynamical consistency of the OTD reduction, arrive at a rather simple *ad hoc* expression for the control force,

$$\mathbf{f}_c = \mathbf{U}\mathbf{Q}\mathrm{diag}[-(\lambda_i + \zeta)\mathscr{H}(\lambda_i)]\mathbf{Q}^\mathsf{T}\mathbf{U}^\mathsf{T}(\mathbf{z} - \mathbf{z}_e), \tag{6}$$

where $\mathscr{H}$ is the Heaviside function, $\zeta \in \mathbb{R}^+$ is a damping parameter, $\mathbf{Q} \in \mathbb{R}^{r\times r}$ is a unitary rotation matrix containing the eigenvectors of $(\mathbf{L}_r + \mathbf{L}_r^\mathsf{T})/2$, and $\{\lambda_i\}_{i=1}^r$ are the eigenvalues of $(\mathbf{L}_r + \mathbf{L}_r^\mathsf{T})/2$ ranked from most (λ_1) to least (λ_r) unstable. The Heaviside function guarantees that the control acts only on directions associated with positive instantaneous growth (those with $\lambda_i \geq 0$), and the parameter ζ governs the intensity with which these directions are damped. The closed-loop rate of change of $\|\boldsymbol{\eta}\|$ is thus negative for all times, thereby ensuring that $\mathbf{z}$ tends to $\mathbf{z}_e$ asymptotically.

Equation (6) assumes that the controller has knowledge of, and can act on, every state variable of the system. In practice, however, the number of sensors and actuators is limited by the apparatus, so we propose a modified OTD control strategy in which the range of the controller is restricted to a small portion of the physical domain. To this end, we consider a subdomain $\bar{\Omega}$ of the original domain Ω, and solve the OTD equations in that subdomain with homogeneous Dirichlet boundary conditions on $\partial\bar{\Omega}$. We denote quantities computed in $\bar{\Omega}$ with an overbar. This allows us to formulate a new control law,

$$\bar{\mathbf{f}}_{c,sub} = \bar{\mathbf{U}}\bar{\mathbf{Q}}\mathrm{diag}[-(\bar{\lambda}_i + \zeta)\mathscr{H}(\bar{\lambda}_i)]\bar{\mathbf{Q}}^\mathsf{T}\bar{\mathbf{U}}^\mathsf{T}\mathbf{R}(\mathbf{z} - \mathbf{z}_e), \tag{7}$$

where $\bar{\mathbf{Q}}$ and $\{\bar{\lambda}_i\}_{i=1}^r$ are the eigenvectors and eigenvalues of $(\bar{\mathbf{L}}_r + \bar{\mathbf{L}}_r^\mathsf{T})/2$, respectively, and $\mathbf{R} \in \mathbb{R}^{m\times d}$ is a rank-m restriction matrix. Here, m denotes the number of degrees of freedom in $\bar{\Omega}$. The presence of the restriction operator left-multiplying the deviation $\mathbf{z} - \mathbf{z}_e$ reflects the fact that only part of the state is accessible. To facilitate application of the control force to the governing equations (which are defined in Ω), we pre-multiply $\bar{\mathbf{f}}_{c,sub}$ by the prolongation matrix $\mathbf{R}^\mathsf{T}$, which gives

$$\mathbf{f}_{c,sub} = \mathbf{R}^\mathsf{T}\bar{\mathbf{U}}\bar{\mathbf{Q}}\mathrm{diag}[-(\bar{\lambda}_i + \zeta)\mathscr{H}(\bar{\lambda}_i)]\bar{\mathbf{Q}}^\mathsf{T}\bar{\mathbf{U}}^\mathsf{T}\mathbf{R}(\mathbf{z} - \mathbf{z}_e). \tag{8}$$

The modified control force (8) is defined in Ω, by construction vanishes outside of $\bar{\Omega}$, and is C^0-continuous across $\partial\bar{\Omega}$.

To select the size and location of the control subdomain, we use a criterion based on the long-time average $\bar{\mu}_j$ of the jth localized instantaneous OTD eigenvalue $\bar{\nu}_j(t) = \langle\bar{\mathbf{L}}\bar{\mathbf{u}}_j, \bar{\mathbf{u}}_j\rangle$. The quantity $\bar{\mu}_j$ may be viewed as a measure for how much information associated with instabilities the localized OTD subspace captures. For example, in flow past a cylinder, we expect $\bar{\mu}_j$ to greatly differ from μ_j if $\bar{\Omega}$ is selected as some region in the far field where the flow is uniform, much more than if $\bar{\Omega}$ includes a substantial fraction of the near field where the wake instability develops and the vortex shedding appears. For a given subdomain size, we found that the optimal subdomain location is where $\bar{\mu}_1$ is the largest.

4.3 Learning the OTD modes from data

For autonomous, invertible, ergodic, and measure-preserving dynamical systems, the OTD modes asymptotically converge to a set of vectors defined at every point on the attractor [5, 3]. In other words, in the post-transient regime, $\mathbf{u}_i$ only depends on the state $\mathbf{z}$, but not on the history of the trajectory, or its own initial condition $\mathbf{u}_i(t_0)$. Hence, we may cease to view $\mathbf{u}_i$ as being parametrized by t, and instead view it as a graph from phase space to tangent space:

$$\begin{aligned}\mathbf{u}_i\colon \mathbb{R}^d &\longrightarrow \mathbb{R}^d\\ \mathbf{z} &\longmapsto \mathbf{u}_i(\mathbf{z}).\end{aligned} \tag{9}$$

In this context, the collection $\{\mathbf{u}_i(\mathbf{z})\}_{i=1}^r$ has been referred to as the "stationary Lyapunov basis" (SLB) at point $\mathbf{z}$ [5].

A promising approach is to learn the mapping (9) from data. This requires several ingredients. First, we assume that we have available a large collection of "snapshots" $\{\mathbf{z}_n\}_{n=1}^d$ for the state. Each $\mathbf{z}_n$ must belong to the attractor, but not necessarily to the same trajectory, a consequence of the use of measure-averaging. Second, we assume that we have a mechanism to compute or reconstruct the vector field $\mathbf{F}(\mathbf{z}_n)$ and the action of the linearized operator $\mathbf{L}$ at $\mathbf{z}_n$ in any direction $\mathbf{v}$. Third, we need to eliminate the explicit dependence of the OTD system (2) on time. This is done by applying the chain rule to the left-hand side of (2), resulting in

$$\nabla_{\mathbf{z}}\mathbf{u}_i\,\mathbf{F}(\mathbf{z}) = \mathbf{L}(\mathbf{z})\mathbf{u}_i - \langle\mathbf{u}_i, \mathbf{L}(\mathbf{z})\mathbf{u}_i\rangle - \sum_{k=1}^{i-1}\left[\langle\mathbf{u}_i, \mathbf{L}(\mathbf{z})\mathbf{u}_k\rangle + \langle\mathbf{u}_k, \mathbf{L}(\mathbf{z})\mathbf{u}_i\rangle\right]\mathbf{u}_k, \quad i \in \{1, \ldots, r\}, \tag{10}$$

where $\nabla_{\mathbf{z}}\mathbf{u}_i$ is the Jacobian of $\mathbf{u}_i$ with respect to $\mathbf{z}$. Although not explicitly shown in (10), the vector $\mathbf{u}_i$ should be understood as $\mathbf{u}_i(\mathbf{z})$.

To learn the graphs $\{\mathbf{u}_i\}_{i=1}^r$ from the collection of snapshots, we employ a neural-network approach. This is appropriate, because each $\mathbf{u}_i$ is a continuous function of $\mathbf{z}$. This allows us to leverage

the universal approximation theorem [6], which states that any function may be approximated by a sufficiently large and deep neural network. We will refer to the learned OTD modes as the "deep OTD modes" (dOTD modes). We assign to each OTD mode its own neural network $\mathbf{u}_i(\mathbf{z};\boldsymbol{\theta}_i)$, where $\boldsymbol{\theta}_i$ denotes the parameters (weights and biases) of the ith network. We use the same fully-connected feed-forward architecture with L hidden layers for all OTD modes. The loss function for the ith network is specified as

$$\ell_i^{pde}(\boldsymbol{\theta}_i) = \frac{1}{N}\sum_{n=1}^{N}\Bigg\| \nabla_{\mathbf{z}}\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i)\,\mathbf{F}(\mathbf{z}_n) - \mathbf{L}(\mathbf{z}_n)\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i) + \langle \mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i), \mathbf{L}(\mathbf{z}_n)\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i)\rangle + \sum_{k=1}^{i-1}\left[\langle \mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i), \mathbf{L}(\mathbf{z}_n)\mathbf{u}_k(\mathbf{z}_n;\boldsymbol{\theta}_i)\rangle + \langle \mathbf{u}_k(\mathbf{z}_n;\boldsymbol{\theta}_i), \mathbf{L}(\mathbf{z}_n)\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i)\rangle\right]\mathbf{u}_k(\mathbf{z}_n;\boldsymbol{\theta}_i)\Bigg\|^2, \tag{11}$$

which is nothing more than the residual of the ith OTD equation in system (10). To ensure that the optimization algorithm converges to an SLB, and not to some other irrelevant minimum, we explicitly enforce orthonormality of the dOTD modes by embedding Gram–Schmidt orthonormalization immediately after the last layer of the network. To ensure that the learning algorithm returns the unique SLB associated with directions of strongest instabilities, we append to the loss function (11) a regularization term, $\ell_i^{lyap}(\boldsymbol{\theta}_i) = -\sigma(\hat{\lambda}_i(\boldsymbol{\theta}_i))$, that penalizes small Lyapunov exponents. Here,

$$\hat{\lambda}_i(\boldsymbol{\theta}_i) = \frac{1}{N}\sum_{n=1}^{N}\langle \mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i), \mathbf{L}(\mathbf{z}_n)\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i)\rangle \tag{12}$$

is the "learned" Lyapunov exponent associated with the ith dOTD mode, and σ is a monotonically increasing, continuous function that exacerbates differences between $\hat{\lambda}_i$'s. Equation (12) is the best approximation of the true exponent λ_i available, given the constraints related to finiteness of the dataset and representability of the OTD modes with neural networks.

The last ingredient to make the method fully data-driven is a mechanism to reconstruct the vector field $\mathbf{F}(\mathbf{z}_n)$ and the action of the Jacobian matrix $\mathbf{L}(\mathbf{z}_n)$ from the collection of snapshots $\{\mathbf{z}_n\}_{n=1}^N$. Reconstruction of $\mathbf{F}(\mathbf{z})$ can be done offline, regardless of the dimensionality of the system. Here, we assume that the snapshots are sampled along a single long trajectory with a uniform and sufficiently small sampling time-step Δt_s, so that we may approximate $\mathbf{F}(\mathbf{z}_n)$ with a standard Euler-forward finite difference. To compute the action of the Jacobian matrix $\mathbf{L}(\mathbf{z})$ from data, we employ the classical algorithm proposed independently by Eckmann et al. [4] and Sano & Sawada [7]. First, we scan the dataset to identify the K nearest neighbors of each datapoint $\mathbf{z}_n$. The nearest neighbors of $\mathbf{z}_n$ are defined as those points $\mathbf{z}_k$ of the orbit (past or future) that are contained in a ball of radius γ centered at $\mathbf{z}_n$:

$$\|\mathbf{z}_n - \mathbf{z}_k\| \le \gamma, \quad k \in \{1,\dots,K\}. \tag{13}$$

If γ is sufficiently small, then each vector $\mathbf{v}_k^n = \mathbf{z}_n - \mathbf{z}_k$ may be viewed as a perturbation vector from the reference orbit. We therefore have

$$\frac{\mathbf{v}_{k+1}^{n+1} - \mathbf{v}_k^n}{\Delta t_s} = \mathbf{L}(\mathbf{z}_n)\mathbf{v}_k^n + \mathcal{O}(\Delta t_s), \tag{14}$$

which allows us to compute the action of the Jacobian matrix $\mathbf{L}(\mathbf{z}_n)$ on the vectors $\mathbf{v}_k^n$. Now, the critical step is to note that the vectors $\mathbf{v}_k^n$ belong to the tangent space at point $\mathbf{z}_n$, and so

do the OTD modes. So if we stack up the vectors $\mathbf{v}_k^n$ into a matrix $\mathbf{V}_n \in \mathbb{R}^{d\times K}$, then the least-square fit $\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i) \approx \mathbf{V}_n\mathbf{V}_n^\dagger\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i)$ should be a reasonably good approximation for the dOTD modes. Here, $\mathbf{V}_n^\dagger$ is the pseudo-inverse of $\mathbf{V}_n$. With this in hand, we can compute the action of the linearized operator on the dOTD modes as

$$\mathbf{L}(\mathbf{z}_n)\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i) \approx \Delta\mathbf{V}_n\mathbf{V}_n^\dagger\mathbf{u}_i(\mathbf{z}_n;\boldsymbol{\theta}_i), \tag{15}$$

where $\Delta\mathbf{V}_n$ is a d-by-K matrix with columns $(\mathbf{v}_{k+1}^{n+1} - \mathbf{v}_k^n)/\Delta t_s$. Equation (15) requires no information other than the snapshot data, and can be used to evaluate the loss function (11).

5 Results

5.1 Stabilization of unsteady flows by OTD control

The great value of the OTD framework has to do with control of instabilities caused by non-normal behavior. Non-normal growth may have severe repercussions on the long-time dynamics, even in cases where modal stability theory predicts asymptotic decay of disturbances. Consider for example plane Poiseuille flow, i.e., pressure-driven flow confined between two rigid, infinitely long, parallel plates. For spanwise wavenumber $\beta = 2$, streamwise wavenumber $\alpha = 0.5$, and Reynolds number $Re = 7000$, linear theory predicts significant non-normal growth of the optimal initial condition (on the order of 1000), followed by asymptotic decay. However, in the full nonlinear problem, sufficiently large non-normal growth triggers transition to turbulence (figure 1a). This outcome cannot be predicted, let along controlled, by classical modal analysis. Yet, figure 1b shows that OTD control is able to suppress non-normal growth, and in turn, transition to turbulence. In contrast, modal control based on the most unstable eigenvector of $\mathbf{L}(\mathbf{w}_e)$ fails at both. This result establishes the superiority of OTD control over modal control in situations dominated by non-normal growth.

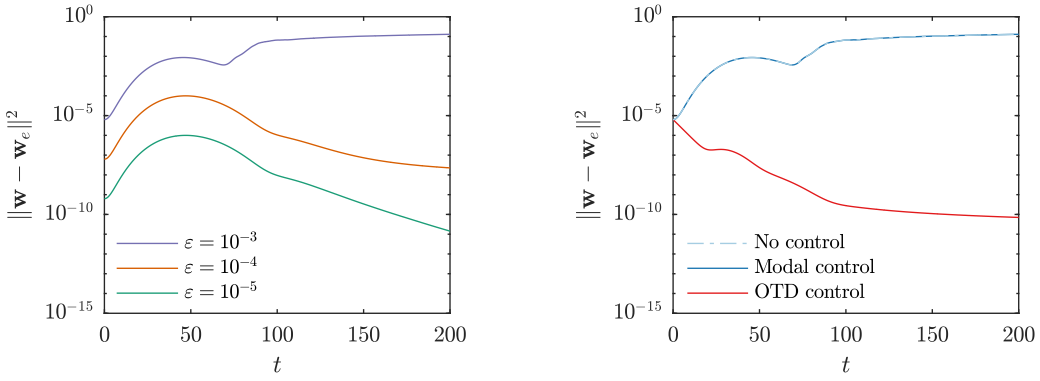
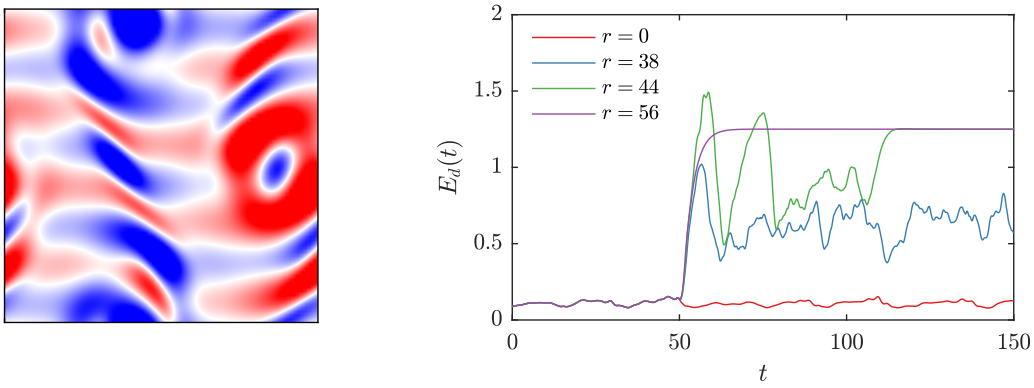


Figure 1: For nonlinear plane Poiseuille flow with $\alpha = 0.5$, $\beta = 2$, and $Re = 7000$, (a) energy of uncontrolled perturbation for various disturbance amplitudes, and (b) for $\varepsilon = 10^{-3}$, energy of perturbation with OTD control ($r = 1$ and $\zeta = 0.1$), modal control based on the most unstable eigenvector of the Orr–Sommerfeld/Squire operator, and no control.

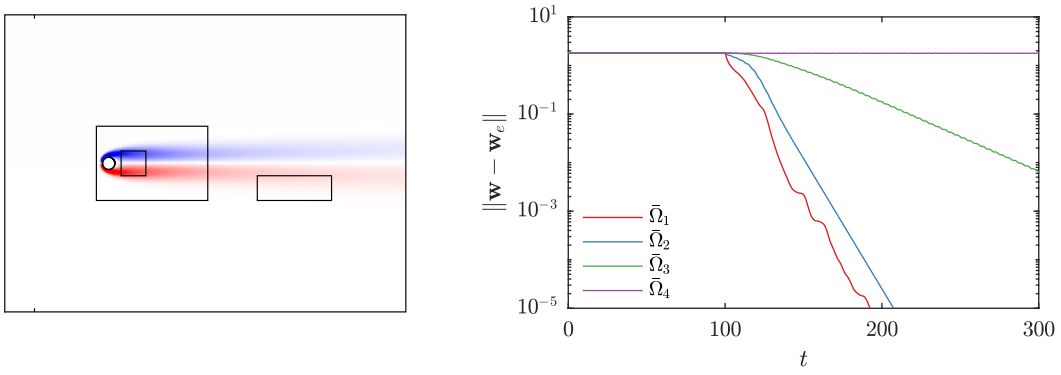
We also found that the OTD control strategy is robust with respect to perturbation amplitude, even in cases in which the trajectory initially evolves on an attractor that lies far away from the target fixed point. Consider for example Kolmogorov flow on the torus $\Omega = [0, 2\pi]^2$ with forcing wavenumber $k_f = 4$ and Reynolds number $Re = 40$, for which the laminar solution $\mathbf{w}_e = (Re/k_f^2)\sin(k_f y)\mathbf{e}_x$ is linearly unstable, and the long-time solution is chaotic (figure 2a). Here, we use OTD control to annihilate the chaotic attractor and steer the trajectory towards $\mathbf{w}_e$. Figure 2b shows time series for the energy dissipation for the uncontrolled case ($r = 0$), and three controlled cases with various values of r. In the absence of control, the trajectory remains on the chaotic attractor and never approaches the laminar solution $\mathbf{w}_e$ (for which $E_d = 1.25$). For $r = 38$, OTD control is not able to annihilate the chaotic attractor. But for $r = 56$, OTD control is able to stabilize $\mathbf{w}_e$. A bissection search shows that no fewer than 44 OTD modes should be included in order to destroy the chaotic attractor. Figure 2b shows that $\mathbf{w}$ approaches $\mathbf{w}_e$ much more rapidly than for $r = 44$, which suggests that including more OTD modes in the controller can prevent "overshoot" of the solution and accelerate stabilization.



(a) (b)

Figure 2: For Kolmogorov flow with $Re = 40$ and $k_f = 4$, (a) spanwise vorticity distribution of the initial condition used in the computations, and (b) energy dissipation for trajectories with OTD control (with $\zeta = 0.1$), and without control. Control is idle in the interval $0 \leq t < 50$ and active for $t \geq 50$.

Finally, we investigate the ability of OTD control to stabilize flow past a cylinder at $Re = 50$ when the spatial range of the controller is limited. We apply OTD control to the subdomains $\{\bar{\Omega}_k\}_{k=1}^4$ shown in figure 3a. We assume that the trajectory initially evolves on the limit cycle, and activate OTD control at $t = 100$. For $r = 8$ and $\zeta = 0.4$, figure 3b shows that the *localized* OTD controller is able to stabilize $\mathbf{w}_e$ *globally* for $\bar{\Omega}_1$, $\bar{\Omega}_2$ and $\bar{\Omega}_3$. These are cases in which the control domain extends over flow regions that are "relevant" to the overall dynamics. By contrast, for the poorly selected domain $\bar{\Omega}_4$, OTD control fails to stabilize $\mathbf{w}_e$. For the three subdomains for which stabilization is achieved, figure 3b shows that the approach to $\mathbf{w}_e$ is faster when the control subdomain is larger. These results suggest that to achieve stabilization, the OTD control subdomain should cover a portion of the computational domain that is relevant to the instability mechanism. Here, "relevance" of a subdomain $\bar{\Omega}$ can be characterized by the leading time-averaged OTD eigenvalue $\bar{\mu}_1$.



(a) (b)

Figure 3: For flow past a cylinder at $Re = 50$, (a) location of control subdomains $\bar{\Omega}_1$ to $\bar{\Omega}_4$ with the spanwise vorticity distribution of $\mathbf{w}_e$ shown in the background, and (b) time series of $\|\mathbf{w}-\mathbf{w}_e\|$ subject to localized OTD control with $r = 8$, $\zeta = 0.4$ for the subdomains shown in (a).

5.2 Learning and predicting extreme events from data

To illustrate the power of the deep OTD modes, we consider a six-dimensional truncation of the classical Charney–DeVore model, with values of the parameters that give rise to significant transitions between "zonal" and "blocked" flow regimes (top panel in figure 4). The intervals of "blocked" flow are considered as extreme events, so we use the neural network approach to learn the mapping $\mathbf{z} \longmapsto \mathbf{u}_i(\mathbf{z})$ in these intervals. We use a neural network with two hidden layers, each with 256 neurons, and 50 data points uniformly sampled over the interval $1000 \leq t \leq 1120$. We reconstruct the Jacobian $\mathbf{L}(\mathbf{z}_n)$ and the velocity field $\mathbf{F}(\mathbf{z}_n)$ from data using the method described earlier. For the Jacobian, we use 60 nearest neighbors for each snapshot. Figure 4 shows that the dOTD modes trained in the interval $1000 \leq t \leq 1120$ can predict subsequent intervals of "blocked" flow. Agreement is poor in intervals of chaos, as expected (neural networks are not always great at extrapolating). What is truly remarkable is that the neural network only needs to know what *one* interval of "blocked" flow looks like to be able to predict all other such intervals in the future.

References

[1] H. Babaee, M. Farazmand, G. Haller, and T. P. Sapsis. Reduced-order description of transient instabilities and computation of finite-time Lyapunov exponents. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27:063103, 2017.

[2] H. Babaee and T. P. Sapsis. A minimization principle for the description of modes associated with finite-time instabilities. *Proceedings of the Royal Society A*, 472:20150779, 2016.

[3] A. Blanchard and T. P. Sapsis. Analytical description of optimally time-dependent modes for reduced-order modeling of transient instabilities. *SIAM Journal on Applied Dynamical Systems*, 18:1143–1162, 2019.

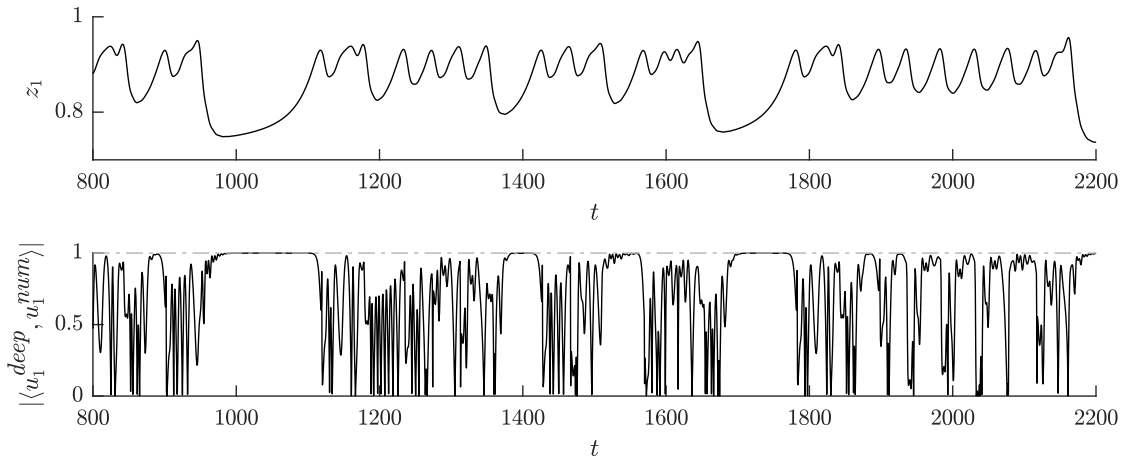


Figure 4: For a six-dimensional truncation of the Charney–DeVore model, time series of the first coordinate z_1 (top), which features several episodes of "blocked" flow; and the absolute value of the inner product between the first OTD mode computed by direct time integration (u_1^{num}) and learned from data by the neural network (u_1^{deep}). The 50 training points are equally spaced in the interval $1000 \leq t \leq 1120$.

[4] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, and S. Ciliberto. Liapunov exponents from time series. *Physical Review A*, 34:4971, 1986.

[5] S. V. Ershov and A. B. Potapov. On the concept of stationary Lyapunov basis. *Physica D: Nonlinear Phenomena*, 118:167–198, 1998.

[6] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2:359–366, 1989.

[7] M. Sano and Y. Sawada. Measurement of the Lyapunov spectrum from a chaotic time series. *Physical Review Letters*, 55:1082, 1985.